

Materialized View Creation for Query Optimization Using Database Client Engine

Ramanand Samdekar

J.D.College of Engineering ,Nagpur
samdekar.ram@gmail.com

Prof. Mirza Moiz Baig

J.D.College of Engineering ,Nagpur
mirzammb@gmail.com

Abstract—Data warehouse and data centric environment is most important area now a days. The researcher required such technologies which reduce time complexity and space complexity. This paper introducing such technology which use cluster base approach to reduce time by clustering table column and table having same size. The common approach of creating materialized view is used ,to reduce time period of accessing data from base table at the same time it will maintain the table as per as requirement of database engine.

Keywords- Database Client Engine ,Cluster Base ,Data Warehouse, Threshold, Materialized Views

I. INTRODUCTION

Data warehouse and its requirement is increasing day by day. Most of the organization also want to store their all data in common data repository and also want to access that data or table of data in minimum time period. Most of the researcher provide technique like creating materialized view of base table on the basis of frequency of execution of query. [1] The proposed work explores the area of a) query clustering for the selection of materialized view to decrease the response time and storage space deployment environment b)Ease network goals c) Enabling data sub-setting d) Enabling disconnected computing. To achieve these benefits, a methodology is proposed in this paper to form a quantitative optimize total query response time under a disk space constraint for data warehouse applications presented in. By clustering common size table also clustering tables having same field is effective methodology to reduce time and space of materialized view.

II. RELATED WORKS

From last two decade the most of researcher are working to reduce the time period or accessing the data from database or data warehouse. The common methodology is use by all is create materialized view of base table. Up till now the materialize view is created by the help of BPSU based on lattice model. By using View adaption technique [2]researcher defined of all SQL clauses and guide line for user and DBA proposed .the advantages of this technique is it reduce cost of query. But disadvantages is only single materialized view changes addressed .View Synchronization method [3]also help to optimization of query time but its disadvantage is only addresses schema change in source and no cost and quality issues are address. View selection SQL based technique[4] is optimal solution for spatial cases (AND/OR view) and polynomial heuristics but problem in AND view graphs not NP-hard .View maintenance method proposed with view selection is more effective but it not properly work with huge amount of data. some time

materialized view take more time as compare to base table due to not proper organization of materialized view. Global view maintenance by maintaining local views in PDMS[5] is One kind of Super peer maintains meta data for mapping schema in inter-peers or inter-peer changes in local PDMS but it has some draw back like Information sharing is complex & difficult in PDMS also Querying not addressed By considering all this factor and research paper we propose system by clustering approach to create materialized view.

III. PROPOSED METHODOLOGY

Queries to DW are critical regarding to their complexity and length[6]. this paper provide cluster based algorithm witch minimize the time of query execution .Initially the practitioner can use random record generator to reduce the time to insert data into the database. without clustering approach is someone want to create materialized view the performance of the data base client engine can be decrease time by time when practitioner data warehouses get increase and also time required to create materialized view is increase. In such case by identifying the table having same length and also identifying the common fields of tables we can cluster the table to generate materialized view. To solve the problem of materialized view selection, a clustering method is proposed in which same queries will be clustered according to their query access frequency to select and generate the materialized views that will minimized the execution time and space complexity .[7] When the query is posed, it will be compared with already clustered or existing query, and the precomputed MV will be returned as a result which will reduce the execution time of the query. In this approach, a framework is created which will reduce the execution time of query when posed to this framework. The proposed frame work as the steps given in the database client view algorithm.

IV. ALGORITHM FOR DATABASE CLIENT VIEW

The steps of the algorithm are as below.

/* we need huge amount of data to compare performance of database client frame work and the base table execution

time. So we use random record generator to insert large amount of data in different tables */

I) Generation of random set of records for given tables in database by record generator.

II) Extraction or generation of all possible set of queries resolved by system on above created records.

/* for materialized view creation and observing the query access time need to measure the query frequency */

III) Optimization of above set of queries according to their access frequency.

IV) Creation of MV according to query access frequency called as Threshold Value and according to Maximum Cluster Area Threshold % According to above step a of MV creation, 3types of MV are created as follows.

1) Single query to Multi table MV. In this response of single query is obtained from multiple MV table.[8]

2) Single query to single table MV. In this response of single query is obtained from single MV table.

3) Multiple queries to single table MV. In this response of multiple similar queries will be obtained from single MV table.

4) After creation of these 3 different types of MV, we will store these MV.

The frame work will also observed that if the query required more time as compare to accessing data from base table at that time the data can be directly access from the base table not from materialized view also if some of the cluster queries are not frequently used for long period that query will automatically deleted by frame work.

V. ALGORITHM MATERIALIZED_VIEW_CREATION

Begin

Step 1.

insert data into different table using Random record generator

/* make sure that the record should be maintain the properties of RDBMS*/

Step 2.

calculate cluster area threshold using following formula:

Table Cluster Area:

Table area (TA) = Ctot*Rtot

Cluster (Ac) = NCRQ*NRRQ

Maximum Cluster Area Threshold %(MCAT) =Ac/TA

Let Threshold_Area_Ratio = TAR

If MCTA <= TAR Then

Create View

Else if Table1_size==Table2_size

create cluster

Else

Ignore

Where, Ctot: Total column of tables

Rtot: Total Records generated for table

Creq: No.of column required by query

Rreq: No.of record required by query

VI. AN ILLUSTRATIVE EXAMPLE

Consider an example Customer record management of a query set where 10 queries are participating and 10 attributes are used in these queries. Say the name of the attributes are A11, A21,.....,A100. [9]

Step 2.

insert record into following table

1)CustomerTBL

2)Product_info

3)Purchase_record

following can be the frequently executed query are as:

1)select all from customerTBL;

2)select * customer where product_id==02;

both query can display same information from base table so that we can cluster the both query and according to it we can generate materialized view.

VII. SELECTION OF VIEWS TO MATERIALIZE DATABASE CLIENT ENGINE

The after clustering query we create materialized data base for table and analyze time of query exaction of base table access ad materialized table access. According the following parameter we can maintain the materialized view[10]

- The query frequency of change

- The size of views

/* if the size of materialized view is large than base table then it cant created.

- The availability of sources: The views, whose data re sides in sources that are rarely available, should be materialized.[11]

- The cost of access: the materialization of views whose data resides in sources with a high cost of access will improve the system performance. Thus, a view will be materialized, if it satisfies at least two criteria.

VIII. CONCLUSIONS

Thus the paper proposes algorithm for the materialized view creation and maitanance.The metodology decrease the data access time from data warehouses.The basic requirement of the any query optimization technique is to craete materalized view on the basis of frequencies of query.the prapose framework provide cluster base approach where frequently exicuted query get cluster and also we can cluster the table having same size so that the the it can reduce space as well as time complexity.

IX. REFERENCES

- [1] An Gong, Weijing Zhao, "Clustering-based Dynamic Materialized View Selection Algorithm"Proceedings of Fifth International Conference on Fuzzy Systems and Knowledge Discovery, 2008, China,pp391- 395.
- [2] A. Gupta, I. Mumick, J. Run, and K. Ross, "Adapting materialized views after redefinitions: techniques and a

- performance study”, In Elsevier Science Ltd., pp 323-362, 2001.
- [3] A. Lee, A. Nica, and E. Rundensteiner, “The EVE approach view synchronization in dynamic distributed environments”, In IEEE Transactions and Data Engineering, 14, 2002.
- [4] A. Gupta, and I. Mumick, “Selection of views to materialize in a data
- [5] warehouse”, In IEEE Transactions on Knowledge and Data Engineering,
- [6] vol. 17, 2005.
- [7] B. Qin, S. Wang, and X. Du, “Effective maintenance of materialized views in peer data management systems”, In Proceedings of First International Conference on Semantics, Knowledge and Grid, 2005.
- [8] S. Agrawal, S. Chaudhari, and V. Narasayya, “Automated selection of
- [9] materialized views and indexes for SQL databases”, In Proceedings of
- [10] 26th International Conference on Very Large Databases, 2000.
- [11] Gupta H, Harinarayan V, Rajaraman A, “Index Selection for OLAP”, Proceeding of International Conference on Data Engineering, 1997, pp. 208-219.
- [12] Mistry H, Roy P, Sudarshan S, et al, “Materialized view selection and maintenance using multi-query optimization”, Proceedings of SIGMOD'01, 2001, pp. 307-3
- [13] P.A.Larsen, Jingren Zhou “Efficient Maintenance of Materialized Outer-Join Views” 23rd International Conference on Data Engineering (ICDE 2007), Istanbul, Turkey, April 2007.
- [14] Xia Sun, Wang Ziqiang “An Efficient Materialized Views Selection Algorithm Based on PSO”. Proceeding of the International Workshop on Intelligent Systems and Applications, ISA 2009, Wuhan, China, May 2009
- [15] S.H. Talebian, S.A. Kareem “Using Genetic Algorithm to Select Materialized Views Subject to Dual Constraints”. International Conference on Signal Processing Systems, Singapore, May 2000